Biochem. J. (1994) 300, 567–575 (Printed in Great Britain)

Novel inducible antibacterial peptides from a hemipteran insect, the sap-sucking bug *Pyrrhocoris apterus*

Stéphane COCIANCICH,* Annick DUPONT,† Gilbert HEGY,† René LANOT,* François HOLDER,* Charles HETRU,* Jules A. HOFFMANN* and Philippe BULET*‡

*Unité Propre de Recherche du CNRS 'Réponse immunitaire et développement chez les insectes', Institut de Biologie Moléculaire et Cellulaire, 15 rue Descartes, 67084 Strasbourg Cédex, and †Laboratoire de Spectrométrie de Masse Bio-Organique de l'Université Louis Pasteur, Unité de Recherche Associée 31 au CNRS 'Chimie Organique des Substances Naturelles', 1 rue Blaise Pascal, 67008 Strasbourg Cédex, France

PD 1994 P: 567-575 (9)

Insects belonging to the recent orders of the endopterygote clade (Lepidoptera, Diptera, Hymenoptera and Coleoptera) respond to bacterial challenge by the rapid and transient synthesis of a battery of potent antibacterial peptides which are secreted into their haemolymph. Here we present the first report on inducible antibacterial molecules in the sap-sucking bug *Pyrrhocoris apterus*, a representative species of the Hemiptera, which predated the Endopterygotes by at least 50 million years in evolution. We have isolated and characterized from immune blood of this species three novel peptides or polypeptides: (*i*) a 43-residue cysteine-rich anti-(Gram-positive bacteria) peptide which is a new member of the family of insect defensins; (*ii*) a 20-residue proline-rich peptide carrying an *O*-glycosylated substitution (*N*-acetylgalactosamine), active against Gram-negative bacteria; (*iii*) a 133-residue glycine-rich polypeptide also active against Gram-negative bacteria. The proline-rich peptide shows high sequence similarities with drosocin, an *O*-glycosylated antibacterial peptide from *Drosophila*, and also with the N-terminal domain of diptericin, an inducible 9 kDa antibacterial peptide from members of the order Diptera, whereas the glycine-rich peptide has similarities with the glycine-rich domain of diptericin. We discuss the evolutionary aspects of these findings.

INTRODUCTION

A series of pioneering studies had established by 1930 that insects can be protected against the injection of normally lethal doses of bacteria by the inoculation of a low dose of bacteria (reviewed in [1]). The induction of this protection, generally referred to as immunization, is a complex and as yet incompletely understood process involving both cellular and humoral components. In higher insects, i.e. essentially in the endopterygote clade, the induced protection is probably to a large extent explained by the rapid synthesis of a battery of potent antibacterial peptides (reviewed in [2] and [3]). This huge clade of insects, which contains more species than the rest of the living world together, is characterized by a far-reaching metamorphosis. The overwhelming species richness in the Endopterygota results from the diversification found in four of its constituent orders: the Coleoptera, the Hymenoptera, the Diptera and the Lepidoptera. Representative species belonging to these four orders have been investigated for the presence of antibacterial peptides, and several families of active inducible molecules have been fully or partially characterized to date (reviewed in [3]). For two of these families, our information is relatively detailed. These are: (*i*) the cecropins (reviewed in [2,4]), which form two α-helices and are active against both Gram-negative and Gram-positive bacteria; (*ii*) the insect defensins, which have six cysteine residues engaged in three intramolecular disulphide bridges ([5,6]; reviewed in [7]). These peptides, which are primarily active against Gram-positive cells, consist of three distinct domains: an N-terminal flexible loop, a central amphipathic α-helix and a C-terminal β-sheet; the α-helix is linked to the β-sheet via two disulphide bridges [8,9]. The other inducible antibacterial peptides of insects have only been characterized at the level of their amino acid sequences (and/or the nucleotide sequences of the corresponding genes). For convenience, and pending more detailed studies, these peptides can be grouped as follows: (*iii*) glycine-rich peptides, namely attacins (Lepidoptera) [10] and the related sarcotoxins II (Diptera) [11], coleoptericin (Coleoptera) [12], diptericin (Diptera) [13] and hymenoptaecin (Hymenoptera) [14]; (*iv*) proline-rich peptides, namely apidaecins (Hymenoptera) [15], abaecin (Hymenoptera) [16] and drosocin (Diptera) [17]. The latter molecules are predominantly active against Gram-negative cells. A remarkable feature of drosocin and diptericin is the presence of *O*-glycosylated substitutions which are necessary for the biological activity of these peptides ([17]; P. Bulet, unpublished work).

The methodologies used to isolate inducible antibacterial peptides from higher insects have failed so far to demonstrate the presence of similar molecules in exopterygote insects, such as members of the Orthoptera and Dictyoptera (lower Neoptera), which have appeared 100 million years before the Endopterygote [18]. Although such insects reportedly build up a protection against lethal doses of bacteria when initially challenged by low doses, they presumably rely on a different protective mechanism possibly on enhancement of phagocytosis.

We are interested in the phylogenetic aspects of the immune response in insects and have addressed in this study the Hemiptera, a major insect group which, although more recent in evolution than the Orthoptera and Dictyoptera, belongs to the exopterygote clade and has not adopted metamorphosis. Earlier studies with two hemipteran species, the sap-sucking bug *Oncopeltus fasciatus* [19] and the blood-sucking bug *Rhodnius prolixus* [20], had shown that a bacterial challenge induces the appearance of antibacterial activity in the haemolymph of these insects.

We now report the isolation from the bug *Pyrrhocoris apterus*

Abbreviations used: m.i.c., minimum inhibitory concentration; c.f.u., colony-forming unit; TFA, trifluoroacetic acid; h.p.g.p.c., high-performance gel permeation chromatography.

of three inducible antibacterial peptides. One of these peptides is a 43-residue new member of the widespread family of insect defensins which are active against Gram-positive bacteria (reviewed in [7]). A second peptide is a proline-rich molecule of 20 amino acids which carries an *O*-glycosylated substitution and is related to drosocin, a peptide which has recently been isolated from the fruitfly, *Drosophila* [17]. This new glycopeptide, which we propose to name 'pyrrhocoricin' (from *Pyrrhocoris apterus*), is active against Gram-negative bacteria. Finally, we have isolated and sequenced a 133-residue polypeptide which is active against Gram-negative bacteria and has some sequence similarity to diptericin, an antibacterial peptide of dipteran insects [13]. It is rich in glycine residues (15%) and we propose the name of hemipteric in (from Hemiptera) for this novel polypeptide.

MATERIALS AND METHODS

Insects, immunization and haemolymph collection

Adults (1000 individuals) of *Pyrrhocoris apterus* (Hemiptera) were collected in the field and received a 2 μl injection containing 2500 cells of *Micrococcus luteus* (Gram-positive strain) and 2500 cells of *Escherichia coli* 1106 (Gram-negative strain). After various time intervals, the haemolymph (about 2 μl per animal) was recovered by sectioning an antenna and gently squeezing the body. The haemolymph was pooled in a precooled polypropylene tube in the presence of the proteinase inhibitor aprotinin (Sigma A-6279; final concn. 10 μg/ml of haemolymph), and of phenylthiourea (final concn. 1 μg/ml of haemolymph) to prevent melanization. The haemolymph was centrifuged at 13000 *g* for 1 h at 4 °C and directly used for the purification of antibacterial peptides.

Bacterial strains and medium

The bacterial strains were gifts from the following colleagues: *E. coli* D31 (streptomycin-resistant), *Serratia marcescens* Db11 and *Enterobacter cloacae* β12 from H. G. Boman (Department of Microbiology, University of Stockholm, Stockholm, Sweden); *E. coli* D22 (an Env A1 mutant with a defection in the outer membrane) from P. L. Boquet (Centre d'Etudes Nucléaires, Saclay, France); *Bacillus megaterium* and *B. subtilis* QB935 from J. Millet and A. Klier (Pasteur Institute, Paris, France); *Pseudomonas aeruginosa* A.T.C.C. 82118, *Alcaligenes faecalis*, *Salmonella typhimurium*, *Staphylococcus aureus*, *Staph. saprophyticus*, *Aerococcus viridans* (= *Gaffkya homarii*), *Listeria monocytogenes*, *Klebsiella pneumoniae* and *Pediococcus acidilactici* from H. Monteil (Institute of Bacteriology, University of Strasbourg, Strasbourg, France); *Micrococcus luteus* A270 and *B. thuringiensis* were from the Pasteur Institute Collection, Paris, France; *E. coli* 1106 was from T. Achstetter (Transgène, Strasbourg, France); *Erwinia carotovora carotovora* (CFBP n° 2141) and *Xanthomonas campestris* pv. *oryzae* (CFBP n° 2532) were from INRA (Angers, France).

All strains were grown on Luria–Bertani's rich nutrient medium [Bactotrypton (1%)/yeast extract (0.5%)/NaCl (1%; w/v)].

Chemicals

Acetonitrile, h.p.l.c. grade, was obtained from Farmitalia Carlo

grade, was obtained from Pierce (Rockford, IL, U.S.A.). Deionized water was produced via a tandem of MilliRO and Milli systems (Millipore).

Antibacterial assays and determination of the minimal inhibitory concentration (m.i.c.)

Antibacterial activity was monitored during the different purification steps by a plate-growth-inhibition assay as described in [6].

The method used for the determination of the m.i.c. has been described in [14]. Briefly, two-fold serial dilutions of *Pyrrhocoris* defensin and pyrrhocoricin were prepared in deionized water and 10 μl aliquots were placed in microtitre plates. Control antibiotic peptides MSI-94 (a broad-spectrum linear amphipathic magainin) and PGLa (a naturally occurring antibiotic peptide from frog) were generously given by M. A. Zasloff (Magainin Pharmaceuticals, Plymouth Meeting, Philadelphia, PA, U.S.A.) The mixture was completed by addition of 100 μl of a bacterial suspension ($A_{600} = 0.001$) in Luria-Bertani's rich nutrient medium. Final concentrations ranged from 0.01 to 5 μM for *Pyrrhocoris* defensin, and from 1 to 10 μM for pyrrhocoricin. The microtitre plates were shaken at 25 °C for 24 h. The m.i.c. values of the antibacterial peptides were expressed according to Casteels et al. [14], as an interval, $a-b$, a being the highest concentration tested at which bacteria are growing and b the lowest concentration that inhibits the growth of the cells.

Bactericidal assay

The methodology used has already been described in [21]: purified anti-(Gram-positive bacteria) peptide (10 μl) was incubated in microtitre plates in the presence of 90 μl of an exponential-phase culture of *M. luteus* (starting $A_{600} = 0.15$) in Luria–Bertani rich nutrient medium. Aliquots were removed at different time intervals and plated on nutrient agar; the number of colony forming units (c.f.u.) was determined after an overnight incubation at 37 °C.

The bactericidal activity of the purified small-sized anti-(Gram-negative bacteria) peptide was determined according to the methodology used for drosocin [17]. Briefly, the purified peptide (10 μl) was incubated in microtitre plates in the presence of 90 μl of an exponential-phase culture of *E. coli* D22 (starting $A_{600} = 0.002$) in phosphate-buffered saline (130 mM). The *E. coli* D22 strain, which has a permeable outer membrane, makes it more sensitive to the antibacterial peptides and therefore suitable when small amounts of antibacterial peptides are available. The activity was detected as described above.

Purification of the antibacterial peptides

Step 1: Sep-Pak pre-purification

The cell-free haemolymph was acidified (0.05% TFA final concn.) and loaded on to a Sep-Pak C_{18} cartridge (Waters Associates). After washing with 5 ml of acidified water (0.05% TFA), stepwise elution was performed with increasing proportions of acetonitrile (20, 50 and 80%) in acidified water. The fractions were concentrated in a vacuum centrifuge (Savant) to remove the organic solvent and TFA and were reconstituted in MilliQ water before monitoring the antibacterial activity by the plate-growth-inhibition assay on *M. luteus* (Gram-positive strain), *E. coli* D22 and *E. coli* D31 (Gram-negative strains).

Step II: size-exclusion chromatography

The first step of purification was a size fractionation by a high-performance gel-permeation chromatography system (h.p.g.p.c.) consisting of serially linked Beckman SEC 3000 and SEC 2000 columns (300 mm × 7.5 mm; Beckman). As the bulk of antibacterial activity was found during the solid-phase extraction on to the Sep-Pak C_{18} cartridge to be present in the 50% elution fraction (see above), only this fraction was applied on the columns and eluted under isocratic conditions with 30% acetonitrile in acidified water at a flow rate of 0.5 ml/min. The column effluent was detected by its u.v. absorption at 225 nm, and the presence of antibacterial activity was monitored as described above.

Step III: final purification

Different conditions were used for the final purification of the active compounds. For the large-sized antibacterial peptide, the active fraction was applied on an Aquapore OD300 C_{18} column (220 mm × 4.6 mm; Brownlee Associates) developed with a linear gradient of 2–52% acetonitrile in acidified water over 90 min at a flow rate of 1 ml/min. The column effluent was monitored by u.v. absorption at 225 nm, and the antibacterial activity was detected as described above. This peptide was further purified by an additional step; the same column was used with a linear gradient of 15–35% acetonitrile in acidified water over 90 min at a flow rate of 2 ml/min. Concerning the small-sized antibacterial peptides, the active fraction was applied to an Aquapore OD300 C_{18} column (220 mm × 4.6 mm, Brownlee Associates). The elution was performed with a linear gradient of 2–52% acetonitrile in acidified water over 90 min at a flow rate of 1 ml/min. The presence of antibacterial activity in the different fractions was detected as described above.

All h.p.l.c. purifications were performed using a Beckman Gold h.p.l.c. system equipped with a Beckman 168 photodiode-array detector.

Enzymic and chemical cleavages

The large-sized antibacterial polypeptide was subjected to digestion with arginyl endopeptidase (Sequencing grade; Takara, Kyoto, Japan), which specifically cleaves the peptide bond on the carboxy side of arginine residues (and also of lysine residues under our conditions), and endoproteinase Glu-C (sequencing grade; Takara), which cleaves the peptide bond on the carboxy side of glutamic acid residues. Digestions were carried out on separate aliquots of 2 nmol of the polypeptide under standard conditions. After a 15 h incubation at 37 °C, the reactions were stopped by adding 20 μl of acidified water. The peptidic fragments were separated on an Aquapore RP300 C_8 column (220 mm × 4.6 mm; Brownlee Associates) developed with a linear gradient of 2–80% acetonitrile in acidified water over 120 min at a flow rate of 0.8 ml/min.

For the chemical cleavage, 5 nmol of the polypeptide were treated with 1000-fold excess of CNBr (Fluka, Buchs, Switzerland) during 20 h at room temperature. The different fragments were separated by chromatography as described above, except for the flow rate, which was 1 ml/min.

peptide was dissolved in 40 μl of 0.5 M Tris/HCl/2 mM EDTA, pH 7.5, containing 6 M guanidinium chloride, to which 2 μl of 2.2 M dithiothreitol were added. The sample was flushed with nitrogen and incubated at 45 °C for 1 h. Freshly distilled 4-vinylpyridine (2 μl) was added and incubated for 10 min at 45 °C in the dark. The pyridylethylated peptide was separated by reversed-phase h.p.l.c. prior to microsequencing.

Microsequence analysis

Automated Edman degradation of the peptides and detection of the phenylthiohydantoin derivatives were performed on a pulse-liquid automatic sequenator (Applied Biosystems; model 473 A).

M.s.

The peptides were dissolved in water/methanol (50:50, v/v) containing 1% acetic acid and analysed on a VG BioTech BioQ mass spectrometer.

Identification of the carbohydrate substitution

The analysis of the carbohydrate was done on 1 nmol of th glycopeptide under the conditions described previously [17 Briefly, the glycopeptide was incubated at 80 °C for 4 h with 1 M HCl in anhydrous methanol and derivatized with *N*-methyl-*N* trimethylsilyltrifluoroacetamide in anhydrous pyridine and tr methylchlorosilane [22]. The component was injected in a Fisor MD800 g.c.–m.s. spectrometer equipped with a JW DB5 fuse silica capillary column (30 m long; 0.32 mm inner diamete 0.1 μm film thickness). The on-column injection's mode w preferred to the Ros one in order to improve the sensitivity the case of very volatile compounds. The sample was injected 50 °C, and the temperature was programmed from 60 to 200 at a rate of 3 °C/min. Helium was used as carrier gas (1.2 m min), and 2–50 ng of material (including the silylating reactive were injected. The mass spectrometer was operated in electro impact mode. M.s. conditions were as follows: electron energ 70 eV; trap current, 150 μA; mass range scanned, from m/z to m/z 650; scan speed, 1 s. The identification of the compon was performed by comparing its retention time and its fr mentogram with that of standard sugars treated according the same procedure.

RESULTS

Isolation and identification of antibacterial peptides

Haemolymph from untreated specimens of *Pyrrhocoris apter* devoid of antibacterial activity. However, the injection of a dose of bacteria induces, within 12–24 h, the appearance strong activity directed against both Gram-positive and Gr negative bacteria. This activity is maintained for several (results not shown).

To isolate and characterize the molecules responsible for activity, we have challenged 1000 adult insects by injection low dose of bacteria and collected their haemolymph on ice 24 h. A total volume of 2.2 ml of haemolymph was obtained after removal of the blood cells by centrifugation, the supern was directly applied to a Sep-Pak C_{18} cartridge for solid-p extraction. The antibacterial activity was recovered by elu with 50% acetonitrile in acidified water. The active fract were pooled and applied to size-exclusion h.p.l.c. col developed under isocratic conditions with 30% acetonitri

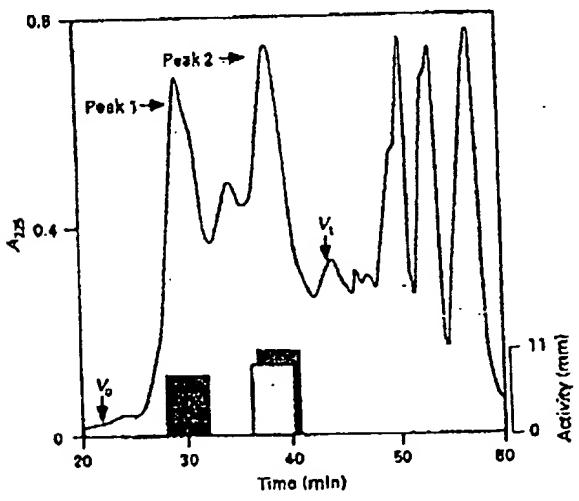

Figure 1 H.p.g.p.c. of immune haemolymph of *P. apterus*

All experimental details were as described in the Materials and methods section. Briefly, active material from the immune haemolymph of 1000 adults of *P. apterus* was pre-purified by solid-phase extraction, and the 50% acetonitrile fraction was analysed on two serial size-exclusion chromatography columns. Antibacterial activity was detected in two fractions (plate-growth-inhibition assay). Peak 1 was active against Gram-negative bacteria (shaded column), whereas peak 2 exhibited both anti-Gram-positive (white column) and -negative activities (shaded column). The antibacterial activity is expressed in diameter (mm) of the growth-inhibition zones. V_0 is the void volume of the column, and V_t is its total volume.

inhibition assay against *M. luteus* and *E. coli* (Figure 1). Anti-(*E. coli*) activity was detected in absorption peak 1 corresponding to compounds with molecular masses ranging from 10 to 30 kDa, and antibacterial activity directed against both *E. coli* and *M. luteus* was observed in absorption peak 2, which contained the molecules with molecular masses lower than 10 kDa. We did not detect antibacterial activities in the other fractions.

The molecules present in peak 1 were subsequently subjected to C_{18} reversed-phase h.p.l.c., eluted with a 2–52% gradient of acetonitrile in acidified water during 90 min at a flow rate of 1 ml/min and re-chromatographed on the same column with a less steep gradient (15–35% of acetonitrile in acidified water over 90 min with an increased flow rate of 2 ml/min). Antibacterial activity was monitored on aliquots of the fractions above. A single anti-(Gram-negative bacteria) compound, appearing pure on the h.p.l.c. chromatogram (results not shown), was recovered and submitted to chemical characterization.

Assuming that the molecule is a polypeptide, we subjected it to Edman degradation (starting material: 240 pmol) and obtained a first sequence of 56 N-terminal residues with an initial yield of 52% and a repetitive yield of 94.5%. We next performed, on separate aliquots of native peptide, enzymic cleavages with arginyl endopeptidase and endoproteinase Glu-C, as well as chemical treatment with CNBr. The fragments obtained were purified by h.p.l.c. and subjected to Edman degradation. The combined sequences (Figure 2a) lead to the conclusion that the peptide is a 133-residue glycine-rich (15%) polypeptide (see also Figure 4 below). The mass calculated from the amino acid sequence (14744.9 Da) is in excellent agreement with the molecular mass measured by m.s. (14743.8 ± 3.5 Da).

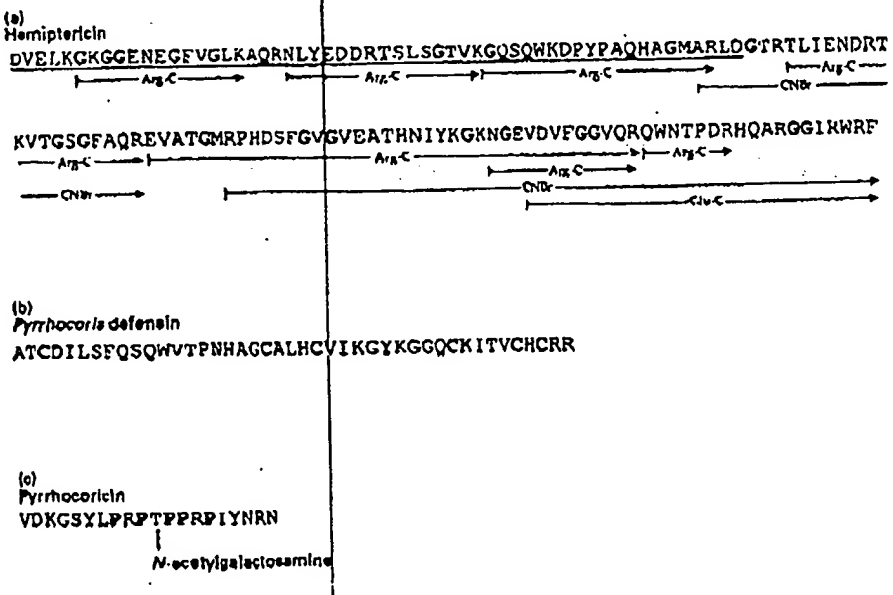

Figure 2 Primary structures of the three antibacterial peptides isolated from the immune haemolymph of *P. apterus*

The residues which are in bold allowed us to classify these molecules into their respective families. (a) Hemiptericin as a glycine-rich peptide. The 56 N-terminal residues obtained after Edman degradation of the native peptide are underlined. The complete amino acid sequence was obtained after enzymic and chemical treatments; arrows indicate the peptidic fragments which were obtained and sequenced after various cleavage treatments of the polypeptide; abbreviations: Arg-C, arginyl endopeptidase; Glu-C, endoproteinase Glu-C. (b) *Pyrrhocoris* defensin as defensin. (c) Pyrrhocoricin as a proline-rich peptide. Pyrrhocoricin is glycosylated on threonine-11 with an *N*-acetylgalactosamine residue.

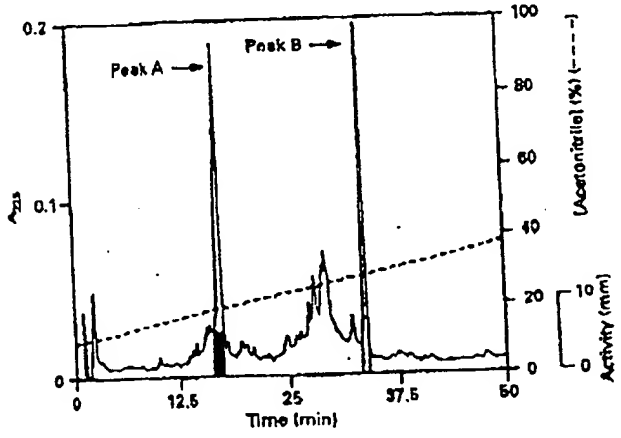

Figure 3 Final purifications of pyrrhocoricin and *Pyrrhocoris* defensin

Peak 2 obtained after the size-exclusion chromatography step was subjected to a reversed-phase h.p.l.c. as described in the Materials and methods section. Two distinct peaks (A and B) contained antibacterial activity. Peak A was active against Gram-negative bacteria (shaded column) and peak B against Gram-positive bacteria (white column). Antibacterial activity is expressed in diameter (mm) of the growth-inhibition zones.

The molecules that were eluted in absorption peak 2 on Figure 1 (corresponding to low-molecular-mass compounds) were subjected to C_{18} reversed-phase h.p.l.c. with a gradient of 2 to 52% of acetonitrile in acidified water during 90 min at a flow rate of 1 ml/min. Two peaks of absorption share an antibacterial activity (Figure 3): peak A was active against *E. coli*, and peak B, which was more hydrophobic, was active against *M. luteus*. The active compounds present in the two peaks appeared sufficiently pure for chemical characterization.

We have first submitted the hydrophobic compound of peak B to Edman degradation and obtained the following sequence of the 15 N-terminal residues: Ala-Thr-Xaa-Asp-Ile-Leu-Ser-Phe-Gln-Ser-Gln-Trp-Val-Thr-Pro. In this sequence, residue 3 could not be assigned. It was immediately apparent that this sequence had a high similarity to that of various members of the family of insect defensins (reviewed in [7]). These peptides have six cysteine residues engaged in three intramolecular disulphide bridges and, assuming that the newly-isolated peptide has similar characteristics, 1 nmol of this molecule was submitted to reduction and alkylation to allow correct identification of the cysteine residues. From 100 pmol of the modified peptide, a sequence of 43 amino acids was subsequently obtained by Edman degradation with an initial yield of 55% and a repetitive yield of 96.1%. This sequence, presented in Figure 2(b) (see also Figure 4), confirms that the peptide is a novel member of the insect defensin family. The calculated mass of the peptide is 4728.5 Da, assuming that the six cysteine residues are engaged in three intramolecular disulphide bridges, as is the case in other defensins (from the fleshflies *Phormia* [23] and *Sarcophaga* [24], and from the dragonfly *Aeschna* [21]). This mass is in excellent agreement with the experimental mass of 4729.3 ± 0.3 Da determined by m.s.

We next analysed the less hydrophobic compound (peak A, Figure 3). Edman degradation on 350 pmol yielded a sequence of 20 amino acids (initial yield of 53%; repetitive yield of 91.5%). Residue 20 gave a signal which strongly suggested that it corresponds to the C-terminal amino acid, as the following cycles gave no signals. From this primary structure, this peptide appears to be particularly rich in proline residues (25%) (Figure 2c; see also Figure 4). The identification of residue 11 was ambiguous,

```
Pyrrhocoris defensin    ATCDILSFQSQWVTPNHAGCALHCVIKGYKGGQCKIT-VCHCRK
Phormia defensin        ATCDLLS----GTGINHSACAAHCLLRGNRGGYCNGKGVCVCRN
Zophobas defensin       FTCDVLGFFIAGTKLNSAACGAHCLALGRRGGYCNSKSVCVCR

Pyrrhocoricin           VDK--GSYLPRPT-PPRPIYNRN
Drosocin                GKPRP-YSPRPTSHPRPI--RV
Apidaecin               GNNRPVYIPQ----PRPPHPRI
Proline-rich domain of dipteric  DEKPK---LILPT--PAPPNLPQ...

Hemipteric              DVELKGKGGENEGFVGLKAQRNLYEDDRTSLSGTVKGQSQWKDPY
Dipteric                                    DEKPKLIL----------PT

                        PAQHAGMARLDGTRTLIENDRTKVTGSGFAQREVATGMRPHDSFG
                        PAPP-------------NLPQLVGGGGG-NRK--------DGFG

                        VGVEATHN-IYKGKNG--EVDVFGGV-QRQWNTP---DRHQAR-G-GIRWRF
                        VSVDA-HQKVWTSDNGRHSIGVTPGYSQH-LGGPYGNSRPDYRIGAGYSYNF
```

Figure 4 Sequence comparison of the antibacterial peptides from *Pyrrhocoris apterus* with antibacterial peptides from the dipteran *Phormia terranovae* (defensin, dipteric) and *Drosophila melanogaster* (drosocin), the hymenopteran *Apis mellifera* (apidaecin) and the coleopteran *Zophobas atratus* (defensin)

Gaps are introduced so that the amino acid sequences of the peptides belonging to each family are maximally aligned. In each case, identical residues and conservative replacements between

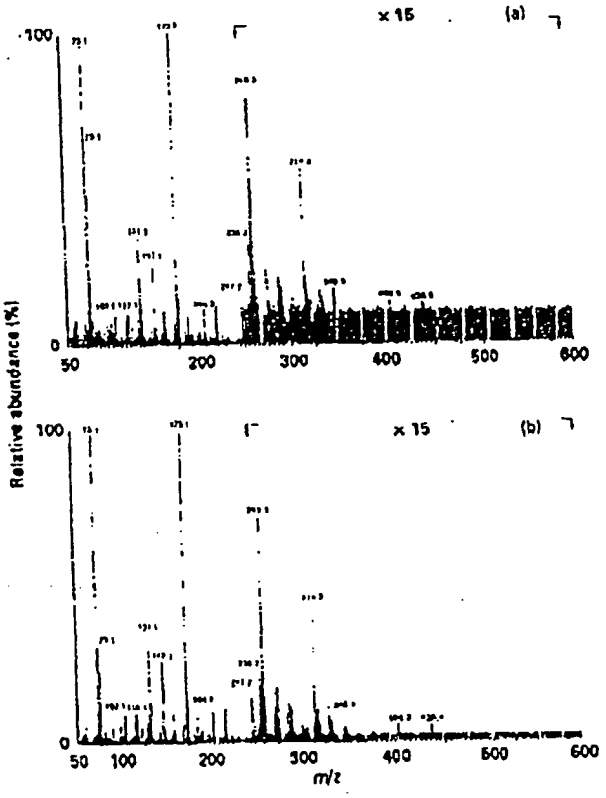

Figure 5 Identification of the carbohydrate substitution of pyrrhocoricin by g.c.–m.s.

Pyrrhocoricin was submitted to methanolysis and pertrimethylsilylation. One major peak appeared in the chromatographic zone corresponding to carbohydrates, whose retention time and fragmentogram were identical with those of the standard *N*-acetylgalactosamine. (a) Electron-impact spectrum of the carbohydrate of pyrrhocoricin; (b) electron-impact spectrum of standard *N*-acetylgalactosamine.

and none of the signals obtained in Edman degradation corresponded to a conventional phenylthiohydantoin amino acid derivative. Interestingly, these signals were found to be identical with those observed in our recent identification and characterization of drosocin, a 19-residue proline-rich inducible antibacterial peptide isolated from *Drosophila*, which has the striking particularity of carrying an *O*-glycosylated substitution on a threonine residue in position 11 [17]. In drosocin, threonine-11 is located within a pseudo-consensus site for *O*-glycosylation [25–27], which is also the case for the residue 11 in this newly isolated peptide. We have consequently assumed, as a working hypothesis, that residue 11 corresponds to a threonine carrying an *O*-glycosylated substitution. The determination of the molecular mass by m.s. gave a value of 2543.3 ± 0.1 Da, which corresponds to the mass calculated for a peptide in which threonine in position 11 is substituted by an *N*-acetylhexosamine residue (2340.3 Da + 203 Da = 2543.3 Da). To confirm the presence of a sugar residue in the native peptide, we next submitted 1 nmol of purified peptide to methanolysis, which cleaves the carbohydrates from the peptide backbone. The resulting mixture was pertrimethylsilylated and analysed by g.c.–m.s. One carbo-

spectrum were found to be identical with those of *N*-acetylgalactosamine (Figure 5).

From the data obtained above we conclude that the proline-rich 20-residue peptide active against *E. coli* carries an *N*-acetylgalactosamine on a threonine residue in position 11 (Figure 2c; see also Figure 4).

We have recently isolated from a new extract of immune haemolymph of *P. apterus* a minor form of this peptide carrying, in addition, a galactose on the *N*-acetylgalactosamine, as is the case in drosocin [17].

Studies on activity spectra and mode of action

Pyrrhocoris defensin

The purified *Pyrrhocoris* defensin was tested by the liquid-growth-inhibition assay against various bacterial strains (Table 1). The peptide has a marked activity (m.i.c. < 2.5 μM) against *M. luteus*, *B. megaterium*, *A. viridans*, *Staph. aureus*, and *Staph saprophyticus*, a moderate activity against *P. acidilactici* and *B. subtilis* QB935 and has no activity against *L. monocytogenes* and *B. thuringiensis*. Except for the D22 form of *E. coli* (shortened lipopolysaccharide), none of the Gram-negative strains tested were affected. Interestingly, it appears that this insect defensin is active against the Gram-positive strains tested in the same range as magainins (MSI-94 and PGLa). These results on the activity spectrum of *Pyrrhocoris* defensin are similar to data obtained with defensin A from *P. terranovae* ([6]; see also [21] for the same conditions of bioassay).

To determine the mode of action of *Pyrrhocoris* defensin, the molecule was tested in the liquid-growth-inhibition assay at a concentration of 1 μM against *M. luteus* (Table 2). Less than 1 min contact with the peptide was found to be sufficient to kill almost all bacteria. This result, indicating that the peptide is bactericidal, is comparable with that observed for the antibacterial activity of insect defensin A from *P. terranovae* [6].

Pyrrhocoricin

The antibacterial spectrum of the purified proline-rich glycosylated peptide from *P. apterus* was also determined using the liquid-growth-inhibition assay (Table 1). Pyrrhocoricin inhibited bacterial growth of the highly sensitive *E. coli* strain D22 at a concentration below 1 μM, which is significantly lower than the m.i.c. values of magainins. The Gram-negative strains *E. coli* 1106, *Ps. aeruginosa* and *Ent. cloacae* β12 were found to be insensitive to the proline-rich peptide of *P. apterus* at an m.i.c. of 5–10 μM. No activity could be detected against the other Gram-negative strains tested. Interestingly, this highly sensitive test allowed us to observe antibacterial activity against two Gram-positive strains: *M. luteus* and *B. megaterium* (m.i.c. 5–10 μM and 1–2.5 μM respectively), in contrast with the less sensitive plate-growth-inhibition assay that we used during the isolation of this peptide (see above).

Finally, as shown in Table 3, purified proline-rich peptide was found to be bactericidal on *E. coli* D22 after a 24 h incubation with the peptide. The kinetics illustrated in Table 3 are similar to those observed with drosocin in *Drosophila* [17]. Note that the kinetics contrast with those of *Pyrrhocoris* defensin, which is bactericidal in less than 1 min of contact with the bacteria (see below).

We were unable to perform similar experiments with the active (Gram-negative bacteria) 133-residue polypeptide isolated in this study, as it was necessary to use all the purified material in

Table 1 Activity spectrum of the insect defensin homologue and of the proline-rich *O*-glycosylated peptide isolated from the immune haemolymph of *P. apterus*

The m.i.c.s. of *Pyrrhocoris* defensin and pyrrhocoricin to inhibit the growth of several representative strains are expressed in final concentrations (μM) and compared with those of control antibiotics (MSI-94, PGLa). Various concentrations of the antibacterial molecules were tested using the liquid growth-inhibition assay described in the Materials and methods section to determine the m.i.c. Abbreviation: n.d., not determined.

Bacterial strain	Gram-positive or Gram-negative	M.i.c. (μM)			
		Pyrrhocoris defensin	MSI-94	PGLa	Pyrrhocoricin
M. luteus	+	0.15–0.3	0.9–1.8	0.3–0.6	5–10
B. megaterium	+	0.15–0.3	0.22–0.45	0.3–0.6	1–2.5
A. viridans	+	0.3–0.6	0.9–1.8	0.6–1.15	n.d.
Staph. aureus	+	0.62–1.25	1.8–3.6	1.15–2.3	> 10
Staph. saprophyticus	+	1.25–2.5	0.9–1.8	0.6–1.15	n.d.
B. subtilis QB935	+	2.5–5	3.6–7.2	4.6–9.2	n.d.
P. acidilactici	+	2.5–5	0.9–1.8	2.3–4.6	n.d.
B. thuringiensis	+	> 5	> 7.2	> 9.2	n.d.
L. monocytogenes	+	> 5	1.8–3.6	1.15–2.3	n.d.
E. coli D22	–	2.5–5	1.8–3.6	1.15–2.3	< 1
E. coli 1106	–	n.d.	1.8–3.6	1.15–2.3	5–10
E. cloacae β12	–	> 5	1.8–3.6	> 9.2	5–10
Ps. aeruginosa A.T.C.C. 82118	–	> 5	1.8–3.6	> 9.2	5–10
S. typhimurium	–	> 5	1.8–3.6	> 9.2	> 10
K. pneumoniae	–	> 5	> 7.2	> 9.2	> 10
S. marcescens Db 11	–	> 5	> 7.2	> 9.2	> 10
A. faecalis	–	> 5	> 7.2	> 9.2	> 10
X. campestris pv *oryzae*	–	n.d.	0.9–1.8	> 9.2	> 10
E. carotovora carotovora	–	n.d.	0.9–1.8	1.15–2.3	> 10

Table 2 Bactericidal effect of *Pyrrhocoris* defensin on *M. luteus*

To 90 μl of an exponential-growth-phase culture of *M. luteus* ($A_{600} = 0.15$) were added 100 pmol of the antibacterial peptide (final concn. 1 μM) or distilled water (control) at zero time. Aliquots were removed at various time intervals and were plated on nutrient agar to determine the number of c.f.u. after an overnight incubation at 37 °C. Results are expressed in 10^6 c.f.u./ml.

Time	C.f.u.	
	1 μM	Control
0 s	39	39
30 s	0	39
0.5 h	0	40
1 h	0	47
3 h	0	65
8 h	0	73
11 h	0	84

Table 3 Bactericidal effect of the proline-rich *O*-glycosylated peptide from *P. apterus* (pyrrhocoricin) on *E. coli* D22

To 90 μl of an exponential-growth-phase culture of *E. coli* D22 ($A_{600} = 0.002$) was added 1 nmol of the pyrrhocoricin (final concn. 10 μM) or distilled water (control) at zero time. Aliquots were removed at various time intervals and were plated on nutrient agar to determine the number of c.f.u. after an overnight incubation at 37 °C. Results are expressed in 10^6 c.f.u./ml.

Time	C.f.u.	
	10 μM	Control
0 s	1.2	1.4
1 min	1.2	1.4
0.5 h	1.2	1.6
2 h	1.2	1.5
6 h	0.6	4.6
24 h	0	16.2

DISCUSSION

Our results provide the first insight into the structures of molecules responsible for the induced antibacterial activity of a hemipteran insect. Three distinct molecules are induced: a novel member of the insect defensin family, a novel proline-rich *O*-glycosylated peptide structurally similar to the recently characterized drosocin of members of the Diptera [17], and a novel glycine-rich polypeptide which shows some sequence similarities with diptericin isolated from dipterans [13]. Interestingly, no cecropins were found in the haemolymph of stimulated *P. apterus*. The novel peptides characterized in bacteria-challenged *P.*

(i) *Pyrrhocoris* defensin

Insect defensins appear as a widespread family of inducible an (Gram-positive bacteria) peptides. They are present in membe of the Diptera, Hymenoptera, Coleoptera and Odonata, but a absent from members of the Lepidoptera (reviewed in [7]). T *Pyrrhocoris* defensin is closest to defensin A of *Phorm terranovae*, which has been characterized at the level of its thre dimensional structure. The *Pyrrhocoris* and *Phormia* defensi have the same relative positions of their six cysteine residu (allowing for a four-residue gap), and 60% of the residues identical, taking into account conservative replacements (Fig has the same thr

dimensional structure as that of *Phormia*, we observed that the major difference between the two peptides resides within the putative N-terminal loop which is longer by four residues in *Pyrrhocoris*. Interestingly, the only other insect order in which defensins have an extended putative N-terminal loop is the Coleoptera (Figure 4), which are the most ancient order within the endopterygote clade. Those two insect defensins have in common 51% of their residues (including conservative replacements).

(ii) The proline-rich *O*-glycosylated peptide pyrrhocoricin

This short novel peptide is interesting in many respects. It is, first of all, extremely hydrophilic: in addition to the polar sugar moiety, 50% of the amino acid residues are charged and/or polar, and only three out of 20 residues are hydrophobic. Of the charged residues, four are basic and one acidic, which explains the strongly cationic character of the peptide (pI 10.3). Pyrrhocoricin shows sequence similarities with drosocin (55%) (Figure 4). They have also in common their short size (20 and 19 residues respectively), the relative richness in proline residues (25 and 33% respectively), their cationic character (pI 10.3 and 12.1 respectively) and the presence at the same position (residue threonine-11) of an *O*-glycosylated substitution (*N*-acetylgalactosamine for pyrrhocoricin and a disaccharide, *N*-acetylgalactosamine-galactose, for drosocin). The presence of minor amounts of a more complex pyrrhocoricin (carrying an *N*-acetylgalactosamine-galactose substitution) underscores the possibility that these substitutions are more complex and have partially been trimmed down during our isolation procedure, which relied on acidic conditions. Pyrrhocoricin shows also some sequence similarities with 18-residue cationic proline-rich antibacterial peptides isolated from honey bees, the apidaecins [15] (Figure 4). A major difference, however, pertains to the absence of *O*-glycosylated substitution in apidaecins, which lack threonine or serine residues. Another common characteristic of these proline-rich antibacterial peptides is the organization of most of their proline residues in proline-arginine-proline triplets (two in pyrrhocoricin, three in drosocin and one in apidaecins).

Interestingly, pyrrhocoricin has some sequence similarity with the N-terminal proline-rich domain of diptericin, a 9 kDa inducible antibacterial peptide from *P. terranovae* [13]. This is illustrated in Figure 4. Unpublished data from this laboratory have recently shown that, in the N-terminal sequence of diptericin, the threonine residue in position 10 also carries an *O*-glycosylated substitution (P. Bulet, unpublished work). Like the other proline-rich short cationic peptides characterized so far, pyrrhocoricin is active against Gram-negative bacteria, but also affects viability of two out of a small number of tested Gram-positive bacterial strains (Table 1). Its exact mode of action remains to be established; as is the case for drosocin, pyrrhocoricin is bactericidal, but its action is relatively slow (Table 3).

(iii) The 133-residue glycine-rich polypeptide hemiptericin

This novel polypeptide is remarkable for the exceptionally large number of charged residues: one amino acid out of three is charged, either positively (24 residues) or negatively (20 residues). As the balance between positively and negatively charged residues is nearly even, the overall charge at pH 7 is only +3.6 (pI 9.4). In addition, the charges appear evenly distributed all along the polypeptide. Computer-assisted analysis has not allowed the

sheet). The molecular-mass determination yielded a valu excellent agreement with the mass calculated on the basis o amino acid sequence, which rules out the existence of substitution in the polypeptide we have isolated.

Hemiptericin shows significant sequence similarity to glycine-rich central and C-terminal domain of diptericin. overall similarity is 29%, and better than 42% if conserv replacements are taken into account (Figure 4). Hemipte shares with the recently isolated 93-residue hymenoptaecin hymenopterans the relative richness in glycine residues an charged residues (one out of three); however, it does not significant sequence similarity to this polypeptide, nor with of the other glycine-rich inducible polypeptides from membe the Lepidoptera (attacins) or Diptera (sarcotoxins II), alth the latter have some sequence similarity with the glycine domain of diptericin [28].

Even if we have only partial information on the ac spectrum and the mode of action of hemiptericin, it is clea the broad-spectrum antibacterial activity in immune ha lymph of the bug *P. apterus* may be the result of: (*i*) indi actions of anti-(Gram-positive bacteria) molecules (insect c sin) and anti-(Gram-negative bacteria) peptide (pyrrhoco and polypeptide (hemiptericin); (*ii*) differences in the kine their mode of actions (slow for pyrrhocoricin, and extreme for *Pyrrhocoris* defensin); (*iii*) the functional relevance of po synergistic mechanisms between these different antibac peptides in the haemolymph. Such a synergism has been sugg to occur between attacin, cecropin and lysozyme in bac challenged silk-moths [29] and between hymenoptaecin daecins and lysozyme in the honey bee [14].

As stated in the Introduction, our information on ind antibacterial peptides in insects is limited to the predom orders of the endopterygote clade (with the exception report on an insect defensin in the order Odonata (dra [21]). These orders evolved after the greatest mass extinc all time, the one between the Permian and Triassic perio million years ago, which wiped out 65% of all insect grou In contrast, the Hemiptera, to which *P. apterus* belong already present in the early Permian. From the phylo point of view, our data show that the facet of the insect im response which is represented by the secretion into the bl induced antibacterial peptides, is relatively ancient wit class of insects. They also show that three distinct cla molecules predominate (minor classes may have escap detection methods): the insect defensins active against positive bacteria, short proline-rich peptides and large g rich polypeptides, the latter being active essentially on negative cells. Insect defensins are present in all the belonging to the endopterygote clade, with the rem exception of the Lepidoptera (reviewed in [7]). The prol peptides, whether glycosylated or not, had so far on detected in two endopterygote orders, the Hymenoptera and the Diptera [17]. They are absent from Lepidopte orders belonging to the Endopterygota contain glyc polypeptides (reviewed in [3]). An interesting result in this is the observation that the proline-rich *O*-glycosylated 20 pyrrhocoricin shows significant sequence similarities with terminal domain of diptericin from Diptera, and that a sequence similarity exists between the glycine-rich 133 hemiptericin and the large glycine-rich domain of th diptericin. This observation lends some credit to the hy that, in higher Diptera, diptericin could have evolved f association of gene domains encoding proline-rich and rich antibacterial peptides [28]. We were unable to de presence of cecropin analogues in *P. apterus*. So far, the

anti-(Gram-positive bacteria) and anti-(Gram-negative bacteria) peptides have only been isolated in Lepidoptera (reviewed in [2]) and Diptera [30], and it may be that these peptides have only appeared after the separation of the dipteran–lepidopteran branch from the endopterygotes in the early Triassic.

We thank Rhône Poulenc Rorer for financial support of G.H. through the program BioAvenir. We also thank Mr. W. Hertzog and his pupils (CM$_2$, Illkirch-Graffenstaden) for their enthusiastic help in collecting several hundred adults of *P. apterus* in the field.

REFERENCES

1 Metalnikow, S. (1927) In Monographie de l'Institut Pasteur (Mascon, et al., eds.). l'Institut Pasteur, Paris
2 Boman, H. G., Faye, I., Gudmunsson, G. H., Lee, J. Y. and Lidholm, D. A. (1991) Eur. J. Biochem. **201**, 23–31
3 Cociancich, S., Bulet, P., Hetru, C. and Hoffmann, J. A. (1994) Parasitol. Today **10**, 132–139
4 Steiner, H., Hultmark, D., Engström, A., Bennich, H. and Boman, H. G. (1981) Nature (London) **292**, 246–248
5 Matsuyama, K. and Natori, S. (1988) J. Biol. Chem. **263**, 17112–17116
6 Lambert, J., Keppi, E., Dimarcq, J. L., Wicker, C., Reichhart, J. M., Dunbar, B., Lepage, P., van Dorsselaer, A., Hoffmann, J. A., Fothergill, J. E. and Hoffmann, D. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 262–266
7 Hoffmann, J. A. and Hetru, C. (1992) Immunol. Today **13**, 411–415
8 Bonmatin, J. M., Bonmat, J. L., Gallet, X., Vovelle, F., Ptak, M., Reichhart, J. M., Hoffmann, J. A., Keppi, E., Legrain, M. and Achstetter, T. (1992) J. Biomol. NMR **2**, 235–256
9 Hanzawa, H., Shimada, I., Kuzuhara, T., Komano, H., Kohda, D., Inagaki, F., Natori, S. and Arata, Y. (1990) FEBS Lett. **269**, 413–420
10 Hultmark, D., Engström, A., Andersson, K., Steiner, H., and Boman, H. G. (1983) EMBO J. **2**, 571–576
11 Ando, K. and Natori, S. (1988) Biochemistry **27**, 1715–1721
12 Bulet, P., Cociancich, S., Dimarcq, J. L., Lambert, J., Reichhart, J. M., Hoffmann, D., Hetru, C. and Hoffmann, J. A. (1991) J. Biol. Chem. **266**, 24520–24525
13 Dimarcq, J. L., Keppi, E., Dunbar, B., Lambert, J., Reichhart, J. M., Hoffmann, D., Rankine, S. M., Fothergill, J. E. and Hoffmann, J. A. (1988) Eur. J. Biochem. **171**, 17–22
14 Casteels, P., Ampe, C., Jacobs, F. and Tempst, P. (1993) J. Biol. Chem. **268**, 7044–7054
15 Casteels, P., Ampe, C., Jacobs, F., Vaeck, M. and Tempst, P. (1989) EMBO J. **8**, 2387–2391
16 Casteels, P., Ampe, C., Rivière, L., Damme, J. V., Elicone, C., Fleming, M., Jacobs, F. and Tempst, P. (1990) Eur. J. Biochem. **187**, 381–386
17 Bulet, P., Dimarcq, J. L., Hetru, C., Lagueux, M., Charlet, M., Hégy, G., van Dorsselaer, A. and Hoffmann, J. A. (1993) J. Biol. Chem. **268**, 14893–14897
18 Lebandeira, C. C. and Sepkoski, J. J., Jr. (1993) Science **251**, 310–315
19 Gingrich, R. E. (1964) J. Insect Physiol. **10**, 179–194
20 Azambuja, P., Freitas, C. C. and Garcia, E. S. (1986) J. Insect Physiol. **32**, 807–812
21 Bulet, P., Cociancich, S., Reuland, M., Sauber, F., Bischoff, R., Hégy, G., van Dorsselaer, A., Hetru, C. and Hoffmann, J. A. (1992) Eur. J. Biochem. **209**, 977–984
22 DeJongh, D. D., Radkurd, T., Hribar, J. D., Hanessian, S., Bieber, M., Dawson, G. and Sweely, C. C. (1969) J. Am. Chem. Soc. **91**, 1728–1740
23 Lepage, P., Bitsch, F., Roecklin, D., Kappi, E., Dimarcq, J. L., Reichhart, J. M., Hoffmann, J. A., Roitsch, C. and van Dorsselaer, A. (1991) Eur. J. Biochem. **196**, 735–742
24 Kuzuhara, T., Nakajima, Y., Matsuyama, K. and Natori, S. (1990) J. Biochem. (Tokyo) **107**, 514–518
25 Wilson, I. B. H., Gavel, Y. and von Heijne, G. (1991) Biochem. J. **275**, 529–534
26 Gooley, A. A., Classon, B. J., Marschalek, R. and Williams, K. L. (1991) Biochem. Biophys. Res. Commun. **178**, 1194–1201
27 O'Connell, B. C., Hagen, F. K. and Tabak, L. A. (1992) J. Biol. Chem. **267**, 25010–25018
28 Wicker, C., Reichhart, J. M., Hoffmann, D., Hultmark, D., Samakovlis, C. and Hoffmann, J. A. (1990) J. Biol. Chem. **265**, 22493–22498
29 Engström, P., Carlsson, A., Engström, A., Tao, Z. and Bennich, H. (1984) EMBO J. **3**, 3347–3351
30 Matsumoto, N., Okada, M., Takahashi, H., Ming, Q. X., Nakajima, Y., Nakanishi, Y., Komano, H. and Natori, S. (1986) Biochem. J. **238**, 717–722